

501P0584 US#

日本国特許庁
PATENT OFFICE
JAPANESE GOVERNMENT

J1017 U.S. PTO
09/834504
04/13/01

別紙添付の書類に記載されている事項は下記の出願書類に記載されて
いる事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed
with this Office.

出願年月日

Date of Application:

2000年 4月14日

出願番号

Application Number:

特願2000-112872

出願人

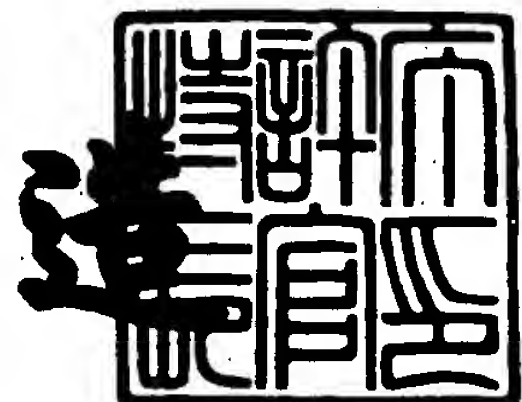
Applicant(s):

ソニー株式会社

2001年 3月 9日

特許庁長官
Commissioner,
Patent Office

及川耕造



CERTIFIED COPY OF
PRIORITY DOCUMENT

出証番号 出証特2001-3016421

【書類名】 特許願

【整理番号】 0000053403

【提出日】 平成12年 4月14日

【あて先】 特許庁長官殿

【国際特許分類】 G10L 5/00

【発明者】

 【住所又は居所】 東京都品川区北品川 6 丁目 7 番 3 5 号 ソニー株式会社
内

 【氏名】 島川 真人

【特許出願人】

 【識別番号】 000002185

 【氏名又は名称】 ソニー株式会社

 【代表者】 出井 伸之

【代理人】

 【識別番号】 100082131

 【弁理士】

 【氏名又は名称】 稲本 義雄

 【電話番号】 03-3369-6479

【手数料の表示】

 【予納台帳番号】 032089

 【納付金額】 21,000円

【提出物件の目録】

 【物件名】 明細書 1

 【物件名】 図面 1

 【物件名】 要約書 1

 【包括委任状番号】 9708842

【ブルーフの要否】 要

【書類名】 明細書

【発明の名称】 情報処理装置および方法、並びに記録媒体

【特許請求の範囲】

【請求項 1】 ユーザにより入力された音声コマンドにより、ネットワークを介して接続された電子機器を制御する情報処理装置において、

前記ネットワークに接続された電子機器を制御するユーザインタフェースコマンドと、前記ユーザにより入力される音声コマンドとの対応関係が定義された音声ユーザインタフェース定義情報を取得する音声ユーザインタフェース定義情報取得手段と、

前記音声ユーザインタフェース定義情報取得手段により取得された前記音声ユーザインタフェース定義情報を記憶する記憶手段と、

前記ユーザにより入力された電子機器を制御する音声コマンドを認識する音声認識手段と、

前記音声認識手段により認識された音声コマンドに対応するユーザインタフェースコマンドを、前記記憶手段により記憶された音声ユーザインタフェース定義情報に基づいて決定する決定手段と、

前記決定手段により決定されたユーザインタフェースコマンドにより、前記電子機器を制御する制御手段と

を備えることを特徴とする情報処理装置。

【請求項 2】 前記音声ユーザインタフェース定義情報は、前記電子機器の ID 情報に対応している

ことを特徴とする請求項 1 に記載の情報処理装置。

【請求項 3】 前記音声ユーザインタフェース定義情報は、前記電子機器の動作状態毎に対応可能な、音声コマンド、ユーザインタフェースコマンド、および、前記ユーザインタフェースコマンドに対応した動作をしたときの動作状態の遷移先の情報を含む

ことを特徴とする請求項 1 に記載の情報処理装置。

【請求項 4】 前記決定手段は、前記電子機器の動作状態を認識する動作状態認識手段をさらに備え、

前記動作状態認識手段により認識した前記電子機器の動作状態に基づいて、前記音声認識手段により認識された音声コマンドに対応するユーザインタフェースコマンドを、前記音声ユーザインタフェース定義情報に基づいて決定することを特徴とする請求項 3 に記載の情報処理装置。

【請求項 5】 前記音声ユーザインタフェース定義情報取得手段は、インターネット上のサーバ、または、記録媒体より前記音声ユーザインタフェース定義情報を取得する

ことを特徴とする請求項 1 に記載の情報処理装置。

【請求項 6】 前記電子機器の ID 情報を取得する ID 情報取得手段をさらに備え、

前記音声ユーザインタフェース情報取得手段は、前記 ID 情報取得手段により取得された前記電子機器の ID 情報に基づいて、対応する前記音声ユーザインタフェース定義情報を取得する

ことを特徴とする請求項 1 に記載の情報処理装置。

【請求項 7】 前記 ID 情報は、ノード ID、機器 ID、機種 ID、または、機器タイプ ID を含む

ことを特徴とする請求項 6 に記載の情報処理装置。

【請求項 8】 前記音声ユーザインタフェース情報取得手段は、前記 ID 情報取得手段により取得された前記電子機器の ID 情報のうち、ノード ID、機器 ID、機種 ID、機器タイプ ID の順位で、対応する前記音声ユーザインタフェース定義情報を取得する

ことを特徴とする請求項 6 に記載の情報処理装置。

【請求項 9】 前記ネットワークに接続された電子機器の接続状態を検出する接続状態検出手段をさらに備え、

前記音声ユーザインタフェース情報取得手段は、前記接続状態検出手段により検出された電子機器の接続状態に基づいて、対応する前記音声ユーザインタフェース定義情報を取得・削除する

ことを特徴とする請求項 1 に記載の情報処理装置。

【請求項 10】 前記ネットワークに接続された電子機器をユーザによって

選択させる選択手段をさらに備え、

前記決定手段は、前記選択手段で選択された機器に対応する前記音声ユーザインタフェース定義情報にのみ基づいて決定を行う機能を有する

ことを特徴とする請求項 1 に記載の情報処理装置。

【請求項 1 1】 前記決定手段により決定されたユーザインタフェースコマンドを履歴として記憶するユーザインタフェースコマンド履歴記憶手段をさらに備え、

前記決定手段により決定されたユーザインタフェースコマンドにより制御すべき、前記ネットワークに接続された電子機器が特定できない場合、ユーザに電子機器を問い合わせ、問い合わせ結果に基づく電子機器を制御するか、または、ユーザインタフェースコマンド履歴記憶手段により記憶された、最近制御された電子機器を制御する

ことを特徴とする請求項 1 に記載の情報処理装置。

【請求項 1 2】 任意に音声ユーザインタフェース定義情報を生成する音声ユーザインタフェース定義情報生成手段をさらに備え、

前記音声ユーザインタフェース定義情報取得手段は、音声ユーザインタフェース定義情報生成手段より前記音声ユーザインタフェース定義情報を取得する

ことを特徴とする請求項 1 に記載の情報処理装置。

【請求項 1 3】 ユーザにより入力された音声コマンドにより、ネットワークを介して接続された電子機器を制御する情報処理装置の情報処理方法において、

前記ネットワークに接続された電子機器を制御するユーザインタフェースコマンドと、前記ユーザにより入力される音声コマンドとの対応関係が定義された音声ユーザインタフェース定義情報を取得する音声ユーザインタフェース定義情報取得ステップと、

前記音声ユーザインタフェース定義情報取得ステップの処理で取得された前記音声ユーザインタフェース定義情報を記憶する記憶ステップと、

前記ユーザにより入力された電子機器を制御する音声コマンドを認識する音声認識ステップと、

前記音声認識ステップの処理で認識された音声コマンドに対応するユーザインタフェースコマンドを、前記記憶ステップの処理で記憶された音声ユーザインタフェース定義情報に基づいて決定する決定ステップと、

前記決定ステップの処理で決定されたユーザインタフェースコマンドにより、前記電子機器を制御する制御ステップと

を含むことを特徴とする情報処理方法。

【請求項 1 4】 ユーザにより入力された音声コマンドにより、ネットワークを介して接続された電子機器を制御する情報処理装置を制御するプログラムであって、

前記ネットワークに接続された電子機器を制御するユーザインタフェースコマンドと、前記ユーザにより入力される音声コマンドとの対応関係が定義された音声ユーザインタフェース定義情報の取得を制御する音声ユーザインタフェース定義情報取得制御ステップと、

前記音声ユーザインタフェース定義情報取得制御ステップの処理で取得された前記音声ユーザインタフェース定義情報の記憶を制御する記憶制御ステップと、

前記ユーザにより入力された電子機器を制御する音声コマンドの認識を制御する音声認識制御ステップと、

前記音声認識制御ステップの処理で認識された音声コマンドに対応するユーザインタフェースコマンドの、前記記憶ステップの処理で記憶された音声ユーザインタフェース定義情報に基づいた決定を制御する決定制御ステップと、

前記決定制御ステップの処理で検索されたユーザインタフェースコマンドにより、前記電子機器を制御する制御ステップと

を含むことを特徴とするコンピュータが読み取り可能なプログラムが記録されている記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、情報処理装置および方法、並びに記録媒体に関し、特に、単一の音声認識装置と統一されたユーザインタフェースにより、複数の電子機器を統合し

て制御できるようにした情報処理装置および方法、並びに記録媒体に関する。

【 0 0 0 2 】

【従来の技術】

音声によりコマンドを送り、電子機器を制御する音声認識技術が普及しつつある。音声認識技術とは、ユーザが発する音声、どのような言葉であるのかを認識する技術であり、例えば、この認識された言葉に対応するコマンドに基づいて、電子機器を制御する装置などに応用されている。

【 0 0 0 3 】

例えば、カーナビゲーション装置などに使用する場合、ユーザが、そのマイクロフォンなどのユーザインタフェース装置に、音声コマンドとして「しんじゅく」を入力すると、カーナビゲーション装置は、マイクロフォンから入力された音声を、内蔵された音声認識装置で解析し、入力された音声「新宿」という言葉であり、これが地名であることを認識する。カーナビゲーション装置は、予め設定されたプログラムにより、認識した地名である「新宿」を目的地としたルート探索を実行し、現在位置から最適なルートを探検して、表示する。

【 0 0 0 4 】

【発明が解決しようとする課題】

しかしながら、上記のような構成では、個々の電子機器に音声認識装置を内蔵させ、個別にユーザからの指示を受付ける形態がとられているため、複数の電子機器を音声により制御する場合、制御したい電子機器の数だけ、音声認識装置を設ける必要があり、現実的な装置を構成することができないという課題があった。

【 0 0 0 5 】

また、この課題を解決する為に、音声認識装置をネットワーク上に設けることが提案されたが、例えば、マイクロフォンなどのユーザインタフェース自体は、各電子機器に設けられる必要があるため、統一的なユーザインタフェースを実現することが困難であるという課題があった。

【 0 0 0 6 】

さらに、新たな電子機器を設けようとすると、対応する音声認識装置およびイ

ンタフェース装置を新たに設けなければならないという課題があった。

【 0 0 0 7 】

さらにまた、上記の構成では、個々の電子機器に内蔵された音声認識装置により、個別にユーザからの指示を受付けるようにされているため、ユーザインタフェースが不統一になると共に、複数機器を統合したコントロールが不可能であるという課題があった。

【 0 0 0 8 】

本発明はこのような状況に鑑みてなされたものであり、単一の音声認識装置と統一されたユーザインタフェースにより、複数の電子機器を統合して制御できるようにするものである。

【 0 0 0 9 】

【課題を解決するための手段】

本発明の情報処理装置は、ネットワークに接続された電子機器を制御するユーザインタフェースコマンドと、ユーザにより入力される音声コマンドとの対応関係が定義された音声ユーザインタフェース定義情報を取得する音声ユーザインタフェース定義情報取得手段と、音声ユーザインタフェース定義情報取得手段により取得された音声ユーザインタフェース定義情報を記憶する記憶手段と、ユーザにより入力された電子機器を制御する音声コマンドを認識する音声認識手段と、音声認識手段により認識された音声コマンドに対応するユーザインタフェースコマンドを、記憶手段により記憶された音声ユーザインタフェース定義情報に基づいて決定する決定手段と、決定手段により決定されたユーザインタフェースコマンドにより、電子機器を制御する制御手段とを備えることを特徴とする。

【 0 0 1 0 】

前記音声ユーザインタフェース定義情報は、電子機器のID情報に対応させるようにすることができる。

【 0 0 1 1 】

前記音声ユーザインタフェース定義情報には、電子機器の動作状態毎に対応可能な、音声コマンド、ユーザインタフェースコマンド、および、ユーザインタフェースコマンドに対応した動作をしたときの動作状態の遷移先の情報を含ませる

ようにすることができる。

【 0 0 1 2 】

前記決定手段には、電子機器の動作状態を認識する動作状態認識手段をさらに設けることができ、動作状態認識手段により認識した電子機器の動作状態に基づいて、音声認識手段により認識された音声コマンドに対応するユーザインタフェースコマンドを、音声ユーザインタフェース定義情報に基づいて決定するようにさせることができる。

【 0 0 1 3 】

前記電子機器のID情報を取得するID情報取得手段をさらに設けることができ、音声ユーザインタフェース定義情報取得手段には、ID情報取得手段により取得された電子機器のID情報に基づいて、対応する音声ユーザインタフェース定義情報を取得させるようにすることができる。

【 0 0 1 4 】

前記ID情報には、ノードID、機器ID、機種ID、または、機器タイプIDを含ませるようにすることができる。

【 0 0 1 5 】

ここでノードIDとはネットワーク上で個々の機器を識別するためのID、機器IDとは製造番号のような機器の個体毎に設けられたID、機種IDとは型名のような同じ機能を持つ機器の集合に対応して設けられたID、機器タイプIDとは例えばTV、ビデオなどのように同種の機能を持つ機器ごとに設定されたIDを示す。

【 0 0 1 6 】

前記音声ユーザインタフェース定義情報取得手段には、ノードID、機器ID、機種ID、機器タイプIDの順位で、機器に対応する音声ユーザインタフェース定義情報の取得を試みさせるようにすることができる。

【 0 0 1 7 】

前記ネットワークに接続された電子機器の接続状態を検出する接続状態検出手段をさらに設けることができ、音声ユーザインタフェース定義情報取得手段には、接続状態検出手段により検出された電子機器の接続状態に基づいて、音声ユーザインタフェース定義情報取得及び破棄を行うようにさせることができる。

【 0 0 1 8 】

前記ネットワークに接続された電子機器を選択する選択手段をさらに設けることができ、決定手段には、選択手段により選択された電子機器に対する音声ユーザインタフェース定義情報のみに基づいて決定させるようにすることができる。

【 0 0 1 9 】

決定手段により検索されたユーザインタフェースコマンドを履歴として記憶するユーザインタフェースコマンド履歴記憶手段をさらに設けることができ、決定手段により決定されたユーザインタフェースコマンドにより制御すべき、ネットワークに接続された電子機器が特定できない場合、ユーザに電子機器を問い合わせ、問い合わせ結果に基づく電子機器を制御するか、または、ユーザインタフェースコマンド履歴記憶手段により記憶された、最近制御された電子機器を制御するようにすることができる。

【 0 0 2 0 】

また、任意に音声ユーザインタフェース定義情報を生成する音声ユーザインタフェース定義情報生成手段をさらに設けることができ、音声ユーザインタフェース定義情報取得手段には、インターネット上のサーバ、記録媒体、または、音声ユーザインタフェース定義情報生成手段より音声ユーザインタフェース定義情報を取得させるようにすることができる。

【 0 0 2 1 】

本発明の情報処理方法は、ネットワークに接続された電子機器を制御するユーザインタフェースコマンドと、ユーザにより入力される音声コマンドとの対応関係が定義された音声ユーザインタフェース定義情報を取得する音声ユーザインタフェース定義情報取得ステップと、音声ユーザインタフェース定義情報取得ステップの処理で取得された音声ユーザインタフェース定義情報を記憶する記憶ステップと、ユーザにより入力された電子機器を制御する音声コマンドを認識する音声認識ステップと、音声認識ステップの処理で認識された音声コマンドに対応するユーザインタフェースコマンドを、記憶ステップの処理で記憶された音声ユーザインタフェース定義情報に基づいて決定する決定ステップと、決定ステップの処理で決定されたユーザインタフェースコマンドにより、電子機器を制御する制

御ステップとを含むことを特徴とする。

【 0 0 2 2 】

本発明の記録媒体のプログラムは、ネットワークに接続された電子機器を制御するユーザインタフェースコマンドと、ユーザにより入力される音声コマンドとの対応関係が定義された音声ユーザインタフェース定義情報の取得を制御する音声ユーザインタフェース定義情報取得制御ステップと、音声ユーザインタフェース定義情報取得制御ステップの処理で取得された音声ユーザインタフェース定義情報の記憶を制御する記憶制御ステップと、ユーザにより入力された電子機器を制御する音声コマンドの認識を制御する音声認識制御ステップと、音声認識制御ステップの処理で認識された音声コマンドに対応するユーザインタフェースコマンドの、記憶ステップの処理で記憶された音声ユーザインタフェース定義情報に基づく決定処理を制御する決定制御ステップと、決定制御ステップの処理で決定されたユーザインタフェースコマンドにより、電子機器を制御する制御ステップとを含むことを特徴とする。

【 0 0 2 3 】

本発明の情報処理装置および方法、並びに記録媒体においては、ネットワークに接続された電子機器を制御するユーザインタフェースコマンドと、ユーザにより入力される音声コマンドとの対応関係が定義された音声ユーザインタフェース定義情報が取得され、取得された音声ユーザインタフェース定義情報が記憶され、ユーザにより入力された電子機器を制御する音声コマンドが認識され、認識された音声コマンドに対応するユーザインタフェースコマンドが、記憶された音声ユーザインタフェース定義情報かに基づき決定され、検索されたユーザインタフェースコマンドにより、電子機器が制御される。

【 0 0 2 4 】

【発明の実施の形態】

図 1 は、本発明に係る音声認識装置 1 の一実施の形態の構成を示す図である。音声認識装置 1 の IEEE (Institute of Electrical and Electronic Engineers) 1394 ネットワークコントローラ 2 1 は、IEEE1394 ネットワーク 3 のインタフェースとして機能し、IEEE1394 ネットワーク 3 に接続されたテレビジョン受像機 (TV

） 4、セットトップボックス（Set Top Box：STB） 5、DVD（Digital Versatile Disk）記録再生装置（DVD） 6、および、DVC（Digital Video Camera） 7と、データを授受する。

【 0 0 2 5 】

IEEE1394ネットワークコントローラ 2 1 には、機器接続検出部 2 1 a が設けられており、IEEE1394ネットワーク 3 上に接続された電子機器の1394Node IDを検出し、音声UI（User Interface）定義管理部 3 2 に出力する。

【 0 0 2 6 】

音声UI定義管理部 3 2 は、利用する音声UI定義情報の追加・削除に関する全体の動作を制御している。また、音声UI定義管理部 3 2 は、機器接続検出部 2 1 a より入力された1394Node IDに基づいて、IEEE1394ネットワーク上に接続された電子機器の接続状態を把握する。さらに、音声UI定義管理部 3 2 は、音声UI定義情報変換部 2 2 に対して、テキスト形式のXML（eXtensible Markup Language）で記述された音声UI定義情報を、IEEE1394ネットワーク 3 上のTV 4、STB 5、DVD 6、DVC 7から取得したり、ネットワークコントローラ 2 7 およびTCP/IPネットワーク（インターネット） 8 を介してWWW（World Wide Web）サーバ 9 から取得したり、ドライブ 2 8 に装着された磁気ディスク 1 0 1、光ディスク 1 0 2、光磁気ディスク 1 0 3、または、半導体メモリ 1 0 4 から取得したり、ROM 2 6、HDD（Hard Disk Drive） 2 9、音声UI定義情報生成部 3 0 から取得するように指示をする。また、このとき音声UI定義管理部 3 2 は同時に機器接続検出部 2 1 a からの電子機器の各種 IDも供給する。音声UI定義情報は、電子機器の動作状態毎に、受付ける音声コマンドと、それに対応するUIコマンドを決定するための情報を記録したものである。尚、XMLについては、後述する。

【 0 0 2 7 】

XMLパーサ 2 2 a は、音声UI定義管理部 3 2 から供給され、XMLで記述された音声UI定義情報を解釈し、その解釈結果を、音声UI定義管理部 3 2 から供給されたID情報と共に、音声認識辞書生成部 2 2 b に出力する。

【 0 0 2 8 】

音声認識辞書生成部 2 2 b は、XMLパーサ 2 2 a から供給された音声UIファイ

ルの解釈結果と、電子機器の1394Node IDに基づいて、音声認識辞書4 1 a, 4 1 b, 4 1 c, 4 1 dを生成し、音声UI定義情報記憶部4 1に出力する。音声認識辞書は、音声UI定義情報に基づいて、各電子機器毎に生成され、その電子機器の動作状態毎に、受付ける音声コマンドと、それに対応するUIコマンドを決定するための情報を記録した内容を、後述する音声認識制御部2 4 aが音声認識の際に利用するフォーマットに変換したものである。尚、音声認識辞書については後述する。

【 0 0 2 9 】

音声認識部2 3の音声UI定義情報記憶部4 1は、音声UI定義管理部3 2により制御され、音声UI定義情報変換部2 2の音声認識辞書生成部2 2 bから供給された音声認識辞書4 1を、対応するそれぞれの電子機器の1394Node IDと共に記憶する。

【 0 0 3 0 】

音声認識制御部2 4 aは音声認識部2 3の全体の動作を制御している。音声認識制御部2 4 aは音声認識辞書を用いて、HMM (Hidden Markov Model : 隠れマルコフモデル) と呼ばれる音声認識手法により、音声入力部2のマイクロフォン (MIC) 2 bから無線通信部2 a, 3 1を介して入力される音声コマンドを認識する。なお、音声認識方法は、HMM以外の方法であってもよい。

【 0 0 3 1 】

さらに、音声認識制御部2 4 aは音声コマンドを送る電子機器が指定された場合については、機器のID情報に基づいて、音声認識辞書として使用する辞書を、対応する音声認識辞書のみを使用するように制限する。

【 0 0 3 2 】

さらにまた、音声認識制御部2 4 aは、認識された音声コマンドの候補となる音声認識結果のうち、上位1 0位までの候補をUIコマンド決定部2 4に出力する。尚、この例においては、音声認識制御部2 4 aは、音声認識結果のうち、上位1 0位までの候補をUIコマンド決定部2 4 bに出力するようにしているが、上位何位までを出力させるかは、任意に設定できるようにしてもよい。

【 0 0 3 3 】

UIコマンド決定部 2 4 b は、音声認識制御部 2 4 a より入力された音声認識結果の上位 1 0 位までの候補を必要に応じてUIコマンド履歴 2 4 c に照合し、最適とされる音声認識結果をUIコマンドとして決定する。UIコマンド決定部 2 4 b は、その決定されたUIコマンドをIEEE1394AV/Cコマンド変換部 2 5 に出力する。

【 0 0 3 4 】

IEEE1394AV/Cコマンド変換部 2 5 は、UIコマンド決定部 2 4 b から入力されたUIコマンドをAV/Cコマンドに変換し、IEEE1394ネットワークコントローラ 2 1 およびIEEE 1394ネットワーク 3 を介して、所定の電子機器に出力する。

【 0 0 3 5 】

ROM 2 6 には、XMLで記述された、各電子機器に共通な、必要最低限のコマンドが記述された音声UI定義情報が、データベースとして記録されている。音声UI定義管理部 3 2 は、他の方法によって対応する音声UI定義情報を取得できないときには、機器接続検出部 2 1 a からのID情報に基づいて、ROM 2 6 に記録された音声UI定義情報を取得するように音声UI定義情報変換部 2 2 に指示を行う。

【 0 0 3 6 】

ネットワークコントローラ 2 7 は、音声UI定義管理部 3 2 の指示に基づいて音声UI定義情報変換部 2 2 によって制御され、TCP/IPネットワーク（インターネット）8 を介して、所定のWWWサーバ 9 にアクセスし、WWWサーバ 9 のHDD 9 a に記録された、電子機器の音声UI定義情報を取得する。

【 0 0 3 7 】

ドライブ 2 8 は、音声UI定義管理部 3 2 の指示に基づいて音声UI定義情報変換部 2 2 によって制御され、装着される磁気ディスク 1 0 1、光ディスク 1 0 2、光磁気ディスク 1 0 3、または、半導体メモリ 1 0 4 に記録された、電子機器の音声UI定義情報を読み出す。HDD 2 9 は、音声UI定義管理部 3 2 の指示に基づいて音声UI定義情報変換部 2 2 によって制御され、記録された音声UI定義情報を読み出す。また、磁気ディスク 1 0 1、光ディスク 1 0 2、光磁気ディスク 1 0 3、半導体メモリ 1 0 4、HDD 2 9 には、IEEE1394ネットワークコントローラ 2 1、または、ネットワークコントローラ 2 7 を介して、音声UI定義情報変換部 2 2 に読み込まれた音声UI定義情報を記録し、再び、必要に応じてそれを読み出

すようにしてもよい。

【 0 0 3 8 】

ユーザ問い合わせ型の音声UI定義情報生成部 3 0 は、上記のいずれの方法でも、音声UI定義情報が入手できない場合、または、ユーザが特に音声UI定義情報を生成できるように設定された場合に使用され、音声入力部 2 の液晶表示部 2 c に音声UI定義情報のエディタを表示させ、ユーザの、MIC 2 b からの入力や選択ボタン 2 d の操作に基づいて音声UI定義情報を生成し、無線通信部 2 a, 3 1 を介して入力された信号に基づいて、音声UI定義情報変換部 2 2 に出力する。

【 0 0 3 9 】

音声認識装置を使用するユーザにより保持される音声入力部 2 では、MIC 2 b から入力される各種の音声コマンドや、選択ボタン 2 d が操作されることにより発生される信号が、無線通信部 2 a を介して、無線で音声認識装置 1 の無線通信部 3 1 に出力される。また、音声入力部 2 の液晶表示部 2 c は、音声UI定義情報を生成するエディタを表示する。ユーザは、この表示に基づいて、音声UI定義情報を編集し、生成することができる。また、音声入力部 2 の液晶表示部 2 c には、音声認識制御部 2 4 a により、特定の電子機器に対するUIコマンドが特定できなかった場合、その候補となる電子機器が表示される。この表示により、ユーザは、選択ボタン 2 d を操作して、UIコマンドの対象となる電子機器を特定することができる。

【 0 0 4 0 】

IEEE1394ネットワーク 3 に接続されたTV 4、STB 5、DVD 6、および、DVC 7 には、CPU (Central Processing Unit)、記憶部 (HDDやメモリなど)、および、通信部 (いずれも図示せず) が設けられており、各電子機器は、そのCPUにより制御され、IEEE1394ネットワーク 3 を介して、TV 4 の記憶部に記憶されている1394Node ID 4 a、機器ID 4 b、機種ID 4 c、または、機器タイプID 4 d、STB 5 の記憶部に記憶されている1394Node ID 5 a、機器ID 5 b、機種ID 5 c、または、機器タイプID 5 d、DVD 6 の記憶部に記憶されている1394Node ID 6 a、機器ID 6 b、機種ID 6 c、機器タイプID 6 d、または、音声UI定義情報 6 e、DVC 7 の記憶部に記憶されている1394Node ID 7 a、機器ID 7 b、機種ID 7 c、または、機

器タイプID 7 d を、それぞれ、(IEEE1394ネットワークに対応した) 通信部から、IEEEネットワーク 3 を介して、音声認識装置 1 のIEEE1394ネットワークコントローラ 2 1 に送信する。

【 0 0 4 1 】

1394Node ID 4 a 乃至 7 a は、IEEE1394ネットワーク 3 上のIDであり、IEEE1394ネットワーク 3 に接続される毎に(バスリセットされる毎に) 付されるIDである。機器IDは、各電子機器を認識する為のIDであり、例えば製造メーカや型式番号、製造番号などの組み合わせによって構成され、その電子機器を一意に決定することを可能とする。この例においては、IEEE1394ネットワーク 3 によりネットワークが構成されているので、各種の処理では、1394Node IDが使用されているが、例えば、機器を接続するネットワークが、IEEE1394ネットワーク 3 以外のもので構成された場合、機器IDを各電子機器を認識するIDとして使用してもよい。

【 0 0 4 2 】

機種IDは、各電子機器の機種を認識する為のIDであり、例えば製造メーカ、型式番号などの組み合わせによって構成される。機器タイプIDは電子機器の種別を大まかに分類するためのIDであり、例えば、その電子機器が、TVであるか、STBであるか、といったことが認識される。

【 0 0 4 3 】

この例においては、DVD 6 にのみ、専用の音声UI定義情報 6 e が保持されており、ネットワークを介してこの音声UI定義情報が直接取得され、その他の機器については上記ID情報に基づいて適切な音声UI定義情報を取得することによって動作するようになっている。

【 0 0 4 4 】

次に、音声UI定義情報について説明する。音声UI定義情報の最も単純な形式は、状態毎に受付可能な音声コマンド、UIコマンド、および、遷移先State(状態)の対応関係を組合せにしたものである。図 2 は、DVC 7 の音声UI定義情報の概念図を示している。例えば、この音声UI定義情報のState 0 の場合、音声コマンドとして受付可能な音声コマンドは、「電源オン」と「電源を入れて」の2つで

ある。これらの音声コマンドが対応するUIコマンドは、いずれも「Power On」である。遷移先Stateに示されるように、DVC7は、「Power On」の処理を実行すると、その状態は、State0からState1に遷移する。この関係が、図3に示されている。図3に示すように、このState0の場合、「電源オン」と「電源を入れて」という音声コマンドに対応した電源を入れるといった処理ができ、状態は、図中の番号1で示すように、State0からState1に遷移する。

【0045】

また、State1の場合、受付可能な音声コマンドは、「電源オフ」、「電源を切って」、および「再生」である。音声コマンドの「電源オフ」、「電源を切って」に対応するUIコマンドは、「Power Off」であり、音声コマンドの「再生」に対応するUIコマンドは、「Play」である。DVC7は、「Power Off」を実行すると、図中の番号2で示すように、State1からState0に遷移し、「Play」を実行すると、図中の番号3で示すように、State1からState2に遷移する。

【0046】

さらに、State2の場合、受付可能な音声コマンドは、「停止」であり、対応するUIコマンドは、「Stop」となり、これが実行されると、状態は、図中の番号4で示すように、State2からState1に遷移する。

【0047】

このように、電子機器の状態により、受付可能な音声コマンドが限定されており、さらに、その音声コマンドにより、その状態も、決まった状態に遷移する。このため、音声認識制御部24aは、音声認識する際に、この状態に基づいて、その状態でしか受け付けられない音声コマンドに関してだけが有効となるように、音声認識処理を行う。

【0048】

次に、この音声UI定義情報を実際に記述するためのXMLベースのフォーマットについて説明する。図2に示すように状態がState0乃至2の場合の音声UI定義情報をXMLで記述したものを以下に示す。

【0049】

1 <音声UI定義>

```

2   <定義クラス>機種</定義クラス>
3   <ID>MDX-001</ID>
4   <Date>1999/12/17</Date>
5   <State数>3</State数>
6   <State NO=0>
7       <コマンド>
8           <音声コマンド>電源オン</音声コマンド>
9           <音声コマンド>電源を入れて</音声コマンド>
10          <UIコマンド>Power On</UIコマンド>
11          <遷移先State>1</遷移先State>
12      </コマンド>
13  </State>
14  <State NO=1>
15      <コマンド>
16          <音声コマンド>電源オフ</音声コマンド>
17          <音声コマンド>電源を切って</音声コマンド>
18          <UIコマンド>Power Off</UIコマンド>
19          <遷移先State>0</遷移先State>
20      </コマンド>
21      <コマンド>
22          <音声コマンド>再生</音声コマンド>
23          <UIコマンド>Play</UIコマンド>
24          <遷移先State>2</遷移先State>
25      </コマンド>
26  </State>
27  <State NO=2>
28      <コマンド>
29          <音声コマンド>停止</音声コマンド>
30          <UIコマンド>Stop</UIコマンド>

```

3 1 <遷移先State>1</遷移先State>

3 2 </コマンド>

3 3 </State>

3 4 </音声UI定義>

【0050】

以上に示すように、XMLは、記述内容を示すタグと呼ばれる「<」と「>」で囲まれた中に、内容を表記しており、各コマンドは、<***>と</***>で囲まれた中に記述される。例えば、第3行目に示すように、<ID>と</ID>で囲まれた中に示される内容は、IDであることを示し、その間に記述された「MDX-001」が、実際のIDである。以下に、各行の記述内容について説明する。

【0051】

第1行目は、第34行目までに記述されている内容が、音声UI定義情報であることを示している。第2行目は、この音声UI定義情報が、機種IDに基づいて定義されているものであることを記述している。第3行目は、この音声UI定義情報が、機種ID「MDX-001」に、対応しているものであることを記述している。

【0052】

第4行目は、この音声UI定義情報が生成された日付が、「1999年12月17日」であることを記述している。第5行目は、この音声UI定義情報で扱う状態数(State数)を示しており、図2に示したように、ここでは、状態数は、State 0乃至2の合計3つの状態を示すことが記述されている。

【0053】

第6行目は、それ以降においては、状態が、State 0である場合に使用されるコマンドについて記述することを示している。第7行目は、それ以降の行で、コマンドについて記述することを示している。第8行目は、State 0の状態で、音声コマンドとして「電源オン」が、受け付けられることを記述している。また、第9行目も、第8行目と同様に、「電源を入れて」が音声コマンドとして受け付けられることを記述している。

【0054】

第10行目は、第8行目と第9行目に記述された「電源オン」、「電源を入れ

て」の音声コマンドが、認識されたとき、対応するUIコマンドとして「Power On」が出力されることが記述されている。第11行目は、「Power On」のコマンドを実行したとき、状態が、State0からState1に遷移することが記述されている。第12行目は、第7行目の「コマンド」の記述の括りを示している。第13行目は、第6行目のState0の場合の記述の括りを示している。

【0055】

第14行目は、状態が、State1の場合について、以下に記述することを示している。第15行目は、以下にコマンドについて記述することを示している。第16行目および第17行目は、音声コマンドとして「電源オフ」、「電源を切って」が、State1の状態を受付けられることを記述している。

【0056】

第18行目は、音声コマンドとして、「電源オフ」、「電源を切って」が、認識されたとき、所定の電子機器に「Power Off」のUIコマンドを出力することが記述されている。第19行目は、「Power Off」のUIコマンドが、実行されたときの状態が、State1からState0に遷移することが記述されている。第20行目は、第15行目の「コマンド」の記述の括りを示している。

【0057】

第21行目は、以下にコマンドの記述があることを示している。第22行目は、状態が、State1の場合、音声コマンドとして「再生」が、受付けられることが記述されている。第23行目は、音声コマンドとして「再生」が、認識されたとき、実行されるUIコマンドとして「Play」が出力されることが記述している。第24行目は、UIコマンド「Play」が、実行されることにより、状態が、State1からState2に遷移することが記述されている。第25行目は、第21行目の「コマンド」の記述の括りを示している。第26行目は、第14行目のState1の記述の括りを示している。

【0058】

第27行目は、以下にState2の状態の場合について、記述されていることを示している。第28行目は、以下にコマンドの記述があることを示している。第29行目は、State2の場合、音声コマンドとして「停止」が受付けられること

を記述している。第 3 0 行目は、音声コマンドとして「停止」を認識したときに、UIコマンドとして「Stop」を出力することを記述している。第 3 1 行目は、UIコマンドとして「Stop」を実行したとき、状態が、State 2 から State 1 に遷移することが記述されている。第 3 2 行目は、第 2 8 行目の「コマンド」の括弧を示している。第 3 3 行目は、第 2 7 行目の State 2 の場合の記述の括弧を示している。第 3 4 行目は、第 1 行目の「音声UI定義」の記述の括弧を示している。

【 0 0 5 9 】

音声UI定義情報は、以上のように記述され、XMLパーサ 2 2 a により認識され、認識結果が音声認識辞書生成部 2 2 b に出力される。音声認識辞書生成部 2 2 b は、この認識結果に基づいて、音声認識制御部 2 4 a が、音声認識の際に使用する認識情報とコマンド情報を含む内部フォーマットに変換し、電子機器毎に音声認識辞書として出力する。従って、音声認識制御部 2 4 a の処理は、この音声UI定義情報に基づいた処理となる。

【 0 0 6 0 】

次に、音声認識辞書について説明する。音声認識辞書は、上述のように電子機器毎に音声UI定義情報から生成され、音声認識制御部に適した内部フォーマットに変換されたものである。図 1 に示すように、音声認識辞書 4 1 a 乃至 4 1 d は、音声UI定義情報記憶部 4 1 に記憶される。例えば、それぞれ音声認識辞書 4 1 a が、TV 4 の機器IDに、音声認識辞書 4 1 b が、STB 5 の機器IDに、音声認識辞書 4 1 c が、DVD 6 の機器IDに、音声認識辞書 4 1 d が、DVC 7 の機器IDにそれぞれ対応していた場合、音声認識制御部 2 4 a は、音声認識辞書 4 1 a 乃至 4 1 d に基づいて認識処理を行う。

【 0 0 6 1 】

次に、図 4 のフローチャートを参照して、音声認識装置 1 が、IEEE1394ネットワーク 3 に接続された電子機器に対応する音声認識辞書を生成する処理について説明する。

【 0 0 6 2 】

ステップ S 1 において、音声UI定義管理部 3 2 は、IEEE1394ネットワークコントローラ 2 1 の機器接続検出部 2 1 a に、IEEE1394ネットワーク 3 に接続が開始

された電子機器の有無を問い合わせる。ステップ S 1 において、機器接続検出部 2 1 a から接続が開始された電子機器がないとの応答があった場合、その処理は、ステップ S 2 に進む。

【 0 0 6 3 】

ステップ S 2 において、音声 UI 定義管理部 3 2 は、機器接続検出部 2 1 a に、接続が解除された電子機器の有無を問い合わせる。ステップ S 2 において、接続が解除された電子機器はないと判定された場合、その処理は、ステップ S 1 の処理に戻り、接続が解除された電子機器があると判定された場合、その処理は、ステップ S 3 に進む。

【 0 0 6 4 】

ステップ S 3 において、音声 UI 定義管理部 3 2 は、接続が解除された電子機器の機器 ID を受け取り、これに基づいて、接続が解除された電子機器の機器 ID に対応する音声認識辞書を音声 UI 定義情報記憶部から削除し（または、使用不可のフラグを設定し）、ステップ S 1 の処理に戻る。

【 0 0 6 5 】

ステップ S 1 において、接続を開始した電子機器があると判定された場合、ステップ S 4 において、音声 UI 定義管理部 3 2 は、IEEE1394 ネットワークコントローラ 2 1 経由で電子機器の音声 UI 定義情報が取得できるか否かを判定する。ステップ S 4 において、IEEE1394 ネットワークコントローラ 2 1 経由で、音声 UI 定義情報が取得できないとの応答があったと判定された場合、その処理は、ステップ S 5 に進む。

【 0 0 6 6 】

ステップ S 5 において、音声 UI 定義管理部 3 2 は、ネットワークコントローラ 2 7 を介して、インターネット 8 上の WWW サーバ 9 から音声 UI 定義情報を取得できるか否かを判定し、取得できないとの応答があったと判定した場合、その処理は、ステップ S 6 に進む。

【 0 0 6 7 】

ステップ S 6 において、音声 UI 定義管理部 3 2 は、ドライブ 2 8 または HDD 2 9 から音声 UI 定義情報が取得できるか否かを判定する。ステップ S 6 において、

ドライブ 2 8 または HDD 2 9 から音声 UI 定義情報が取得できないと判定された場合、その処理は、ステップ S 7 の処理に進む。

【 0 0 6 8 】

ステップ S 7 において、音声 UI 定義管理部 3 2 は、ユーザ問い合わせによる音声 UI 定義情報生成を行う設定になっているか否かを判定する。ステップ S 7 において、ユーザ問い合わせによる音声 UI 定義情報生成を行う設定になっていないと判定された場合、その処理は、ステップ S 8 に進む。

【 0 0 6 9 】

ステップ S 8 において、音声 UI 定義管理部 3 2 は、ROM 2 6 に問い合わせ、入手した TV の 1394Node ID に対応する必要最低限の音声 UI 定義情報を取得し、XML パーサ 2 2 a に出力する。ステップ S 9 において、XML パーサ 2 2 a は、入力された音声 UI 定義情報を解析し、その解析結果を音声認識辞書生成部 2 2 b に出力する。音声認識辞書生成部 2 2 b は、入力された解析結果から、音声認識辞書を生成し、1394Node ID と共に音声認識部 2 3 の音声認識制御部 2 4 a に出力する。音声認識制御部 2 4 a は、入力された音声認識辞書を音声 UI 定義情報記憶部 4 1 に出力し、ステップ S 1 の処理に戻りそれ以降の処理が繰り返される。

【 0 0 7 0 】

ステップ S 4 において、例えば、DVD 6 が接続開始された場合、DVD 6 は、音声 UI 定義情報 6 e を備えているので、音声 UI 定義管理部 3 2 は、IEEE1394 ネットワークコントローラ 2 1 から音声 UI 定義情報が取得可能であると判定し、その処理は、ステップ S 1 1 に進む。

【 0 0 7 1 】

ステップ S 1 1 において、音声 UI 定義管理部 3 2 は、IEEE1394 ネットワークコントローラ 2 1 を介して DVD 6 の音声 UI 定義情報 6 e を取得し、XML パーサ 2 2 a に出力し、その処理は、ステップ S 9 に進む。

【 0 0 7 2 】

ステップ S 5 において、ネットワーク上の WWW サーバ 9 から音声 UI 定義情報を取得できると判定された場合、その処理は、ステップ S 1 2 に進む。

【 0 0 7 3 】

ステップ S 1 2 において、音声 UI 定義管理部 3 2 は、ネットワークコントローラ 2 7 を制御し、TCP/IP ネットワーク（インターネット）8 を介して、WWW サーバ 9 にアクセスし、その HDD 9 a に記録された音声 UI 定義情報を取得し、XML パーサ 2 2 a に出力し、その処理は、ステップ S 9 に進む。

【 0 0 7 4 】

ステップ S 6 において、ドライブ 2 8 または HDD 2 9 から音声 UI 定義情報が取得できると判定された場合、その処理は、ステップ S 1 3 に進む。

【 0 0 7 5 】

ステップ S 1 3 において、音声 UI 定義管理部 3 2 は、ドライブ 2 8 に装着されている磁気ディスク 1 0 1、光ディスク 1 0 2、光磁気ディスク 1 0 3、若しくは、半導体メモリ 1 0 4、または、HDD 2 9 にアクセスし、音声 UI 定義情報を取得し、XML パーサ 2 2 a に出力し、その処理は、ステップ S 9 に進む。

【 0 0 7 6 】

ステップ S 7 において、ユーザ問い合わせによる音声 UI 定義情報を生成させる設定になっている場合、その処理は、ステップ S 1 4 に進む。

【 0 0 7 7 】

ステップ S 1 4 において、音声 UI 定義管理部 3 2 は、ユーザ問い合わせ型の音声 UI 定義情報生成部 3 0 に問い合わせ、音声 UI 定義情報を生成させる。このとき、音声 UI 定義情報生成部 3 0 は、無線通信部 3 1、2 a を介して音声入力部 2 に問い合わせ、音声 UI 定義情報を生成する為のエディタを液晶表示部 2 c に表示させる。このとき、例えば、図 2 に示されていたような音声 UI 定義情報を生成する場合、図 5 に示すように、液晶表示部 2 c にはエディタが表示され、ユーザは、各状態と遷移先の関係を音声コマンドと UI コマンドを記述することで音声 UI 定義情報を生成する。

【 0 0 7 8 】

すなわち、図 5 に示すように、状態（State）を描き、各種 State 間を結ぶ矢印により状態遷移を示す。その際、矢印の傍らに、各状態で受付可能な音声コマンドと UI コマンドを併記することで、音声 UI 定義情報を生成する。この例に、おいては、この音声 UI 定義情報の State 0 の場合、音声コマンドとして受付可能な音

声コマンドは、「電源オン」と「電源を入れて」の2つである。これらの音声コマンドが対応するUIコマンドは、いずれも「Power On」である。遷移先Stateに示されるように、DVC 7は、「Power On」の処理を実行すると、その状態は、State 0からState 1に遷移する。

【 0 0 7 9 】

また、State 1の場合、受付可能な音声コマンドは、「電源オフ」、「電源を切って」、および「再生」である。音声コマンドの「電源オフ」、「電源を切って」に対応するUIコマンドは、「Power Off」であり、音声コマンドの「再生」に対応するUIコマンドは、「Play」である。DVC 7は、「Power Off」を実行すると、State 1からState 0に遷移し、「Play」を実行すると、State 1からState 2に遷移する。

【 0 0 8 0 】

さらに、State 2の場合、受付可能な音声コマンドは、「停止」であり、対応するUIコマンドは、「Stop」となり、これが実行されると、状態は、図中の番号4で示すように、State 2からState 1に遷移する。

【 0 0 8 1 】

その後、生成された音声UI定義情報は、無線通信部 2 a, 3 1 および音声定義ファイル生成部 3 0 を介して音声UI定義管理部 3 2 に出力される。そこで、音声UI定義管理部 3 2 は、取得した音声UI定義情報をXMLパーサ 2 2 a に出力し、その処理は、ステップ S 9 に進む。

【 0 0 8 2 】

次に、図 6 のフローチャートを参照して、音声認識処理について説明する。ここで、音声認識装置 1 上では、図 4 のフローチャートの処理により、既に、IEEE 1394 ネットワーク 3 上に接続された各電子機器の音声認識辞書が、音声UI定義情報記憶部 4 1 に記憶されているものとする。

【 0 0 8 3 】

音声入力部 2 のMIC 2 b からユーザにより、音声コマンドが入力されると処理が開始される。ステップ S 2 1 において、音声認識制御部 2 4 a は、無線通信部 3 1 より入力された音声コマンドを、音声UI定義情報記憶部 4 1 に記憶された

音声認識辞書 4 1 a 乃至 d に基づいて認識処理を行い、所定のコマンドとみなされる上位 1 0 位までの UI コマンド候補を UI コマンド決定部 2 4 b に出力する。

【 0 0 8 4 】

ステップ S 2 2 において、UI コマンド決定部 2 4 b は、対象電子機器と UI コマンドを決定できたか否かを判定し、対象電子機器と UI コマンドを決定することができたと判定した場合、その処理は、ステップ S 2 3 に進む。

【 0 0 8 5 】

ステップ S 2 3 において、UI コマンド決定部 2 4 b は、所定の電子機器の UI コマンドを IEEE1394AV/C コマンド変換部 2 5 に出力する。ステップ S 2 4 において、IEEE1394AV/C コマンド変換部 2 5 は、入力された UI コマンドを、AV/C コマンドに変換し、IEEE1394 ネットワーク 3 を介して、所定の電子機器に出力する。

【 0 0 8 6 】

ステップ S 2 2 において、対象電子機器と UI コマンドが決定できなかったと判定された場合、ステップ S 2 5 において、UI コマンド決定部 2 4 b は、UI コマンドが決定できたか否かを判定し、UI コマンドが決定できたと判定した場合、すなわち、対象電子機器が、複数になってしまったような場合、その処理は、ステップ S 2 6 に進む。

【 0 0 8 7 】

ステップ S 2 6 において、UI コマンド決定部 2 4 b は、候補となる電子機器を出力する。ステップ S 2 7 において、UI コマンド決定部 2 4 b は、無線通信部 3 1, 2 a を介して、音声入力部 2 の液晶表示部 2 c に候補となる電子機器として、例えば、DVD 6 と DVC 7 が候補となる電子機器としてあげられた場合、図 7 に示すように、音声入力部 2 の液晶表示部 2 c に表示される。ユーザは、この表示に基づいて、選択ボタン 2 d を操作して、「DVD」か「DVC」のいずれかを選択する。この選択ボタンは、図 7 に示すようなボタンでもよいし、また、タブレット式のボタンでもよい。さらに、液晶表示部 2 c をタッチパネルとして、液晶表示部を触れることで決定するようにしてもよい。そして、UI コマンド決定部 2 4 b は、候補の電子機器がいずれであるかについての選択結果を受信した後、ステップ S 2 3 以降の処理が繰り返される。

【 0 0 8 8 】

ステップ S 2 5 において、コマンドが決定できなかったと判定された場合、すなわち、全く認識できなかった場合、ステップ S 2 8 において、音声認識制御部 2 4 a は、再入力を求める表示を音声入力部 2 の液晶表示部 2 c に表示させ、再び、ステップ S 1 以降の処理を繰り返す。

【 0 0 8 9 】

ステップ S 2 5 において、UI コマンドが決定したと判定された場合、UI コマンド決定部 2 4 b は、UI コマンド履歴を用いて、対象機器を最近使用した電子機器とするようにしてもよい。

【 0 0 9 0 】

以上においては、IEEE1394AV/C コマンド変換部 2 5 は、AV/C コマンドを生成し、出力するのみであるが、例えば、電子機器から動作完了を示す AV/C レスポンスを受信するようにし、さらに、音声認識制御部 2 4 a に出力するようにしてもよい。

【 0 0 9 1 】

また、以上においては、電子機器を接続するネットワークとして、IEEE1394 を用いた例について説明したが、ネットワークは、これ以外のものを使用してもよく、その際、1394Node ID が使用できないことになるが、機器 ID を、1394Node ID と同様に使用したり、あるいはそのネットワーク上の ID を 1394NodeID に代えて利用すれば、ネットワーク上に、接続された電子機器を識別することができる。

【 0 0 9 2 】

さらに、以上においては、各種 ID は、電子機器に予め付されたものとして説明してきたが、例えば、音声 UI 定義管理部 3 2 が、接続された電子機器に対して、ID を付するようによい。また、機器 ID、機種 ID、および、機器タイプ ID のいずれも、保持していない電子機器を使用する場合、ユーザが音声 UI 定義情報を生成した後、音声 UI 定義管理部 3 2 が、ID を付するようにすることで、上記と同様の操作が可能となる。

【 0 0 9 3 】

さらにまた、音声認識辞書を記憶する音声 UI 定義情報記憶部 4 1 が、必要に

応じて、音声認識辞書を削除する場合、キャッシュメモリをさらに設けて、そこに、移動させることで、一度は削除された音声認識辞書が必要となったときに、そのキャッシュメモリから読み込むようにすれば、音声認識辞書の生成処理を省くことができ、その処理を高速化することができる。

【 0 0 9 4 】

以上によれば、複数の電子機器を単一の音声認識装置とユーザインタフェースにより制御できるようにすることが可能となる。

【 0 0 9 5 】

上述した一連の処理は、ハードウェアにより実行させることもできるが、ソフトウェアにより実行させることもできる。一連の処理をソフトウェアにより実行させる場合には、そのソフトウェアを構成するプログラムが、専用のハードウェアに組み込まれているコンピュータ、または、各種のプログラムをインストールすることで、各種の機能を実行させることが可能な、例えば汎用のパーソナルコンピュータなどに記録媒体からインストールされる。

【 0 0 9 6 】

この記録媒体は、図 1 に示すように音声認識装置 1 に予め組み込まれた状態でユーザに提供される、プログラムが記録されている HDD 2 9 だけではなく、コンピュータとは別に、ユーザにプログラムを提供するために配布される、プログラムが記録されている磁気ディスク 1 0 1 (フロッピーディスクを含む)、光ディスク 1 0 2 (CD-ROM(Compact Disk-Read Only Memory), DVD (Digital Versatile Disk) を含む)、光磁気ディスク 1 0 3 (MD (Mini-Disk) を含む)、もしくは半導体メモリ 1 0 4 (Memory Stickを含む) などよりなるパッケージメディアにより構成される。

【 0 0 9 7 】

尚、本明細書において、記録媒体に記録されるプログラムを記述するステップは、記載された順序に沿って時系列的に行われる処理は、もちろん、必ずしも時系列的に処理されなくとも、並列的あるいは個別に実行される処理を含むものである。

【 0 0 9 8 】

【発明の効果】

本発明の情報処理装置および方法、並びに記録媒体によれば、ネットワークに接続された電子機器を制御するユーザインタフェースコマンドと、ユーザにより入力される音声コマンドとの対応関係が定義された音声ユーザインタフェース定義情報を取得し、取得した音声ユーザインタフェース定義情報を記憶し、ユーザにより入力した電子機器を制御する音声コマンドが認識し、認識した音声コマンドに対応するユーザインタフェースコマンドを、記憶した音声ユーザインタフェース定義情報から検索し、検索したユーザインタフェースコマンドにより、電子機器を制御するようにしたので、単一の音声認識装置と統一されたユーザインタフェースにより、複数の電子機器を統合して制御することが可能となる。

【図面の簡単な説明】

【図 1】

本発明を適用した音声認識装置のブロック図である。

【図 2】

音声UI定義情報を説明する図である。

【図 3】

状態の遷移を説明する図である。

【図 4】

音声認識辞書を生成する処理を説明するフローチャートである。

【図 5】

音声UI定義情報を生成するときのエディタの表示例である。

【図 6】

音声認識処理を説明するフローチャートである。

【図 7】

UIコマンドの対象となる電子機器の候補を選択するときの表示例である。

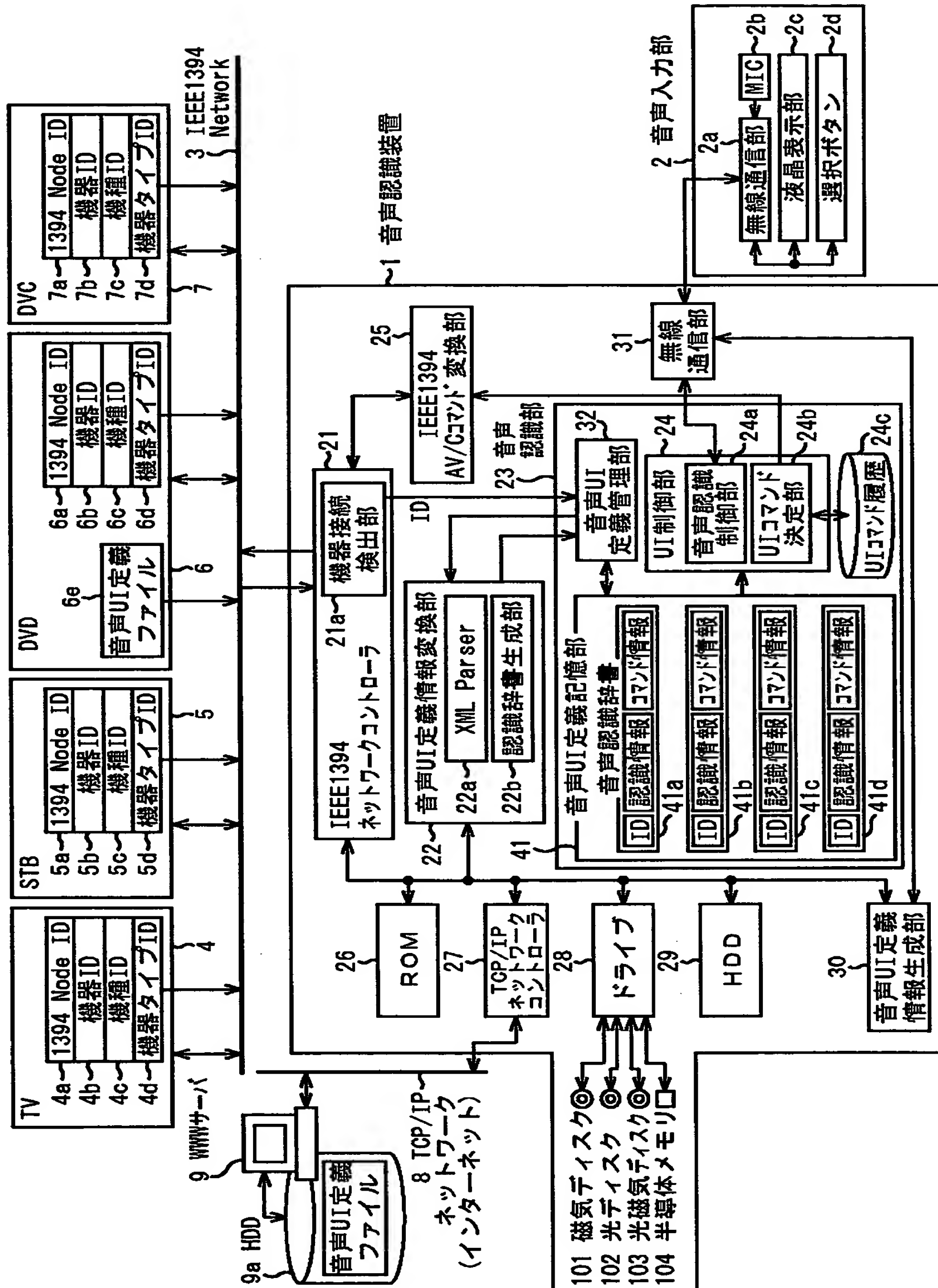
【符号の説明】

1 音声認識装置, 2 音声入力部, 2 a 無線通信部, 2 b MIC, 2 c 液晶表示部, 2 d 選択ボタン, 3 IEEE1394ネットワーク, 4 TV, 4 a 1394Node ID, 4 b 機器ID, 4 c 機種ID, 4 d 機器タイプID, 5 STB, 5 a

1394Node ID, 5 b 機器ID, 5 c 機種ID, 5 d 機器タイプID, 6 DVD,
6 a 1394Node ID, 6 b 機器ID, 6 c 機種ID, 6 d 機器タイプID, 7 D
VC, 7 a 1394Node ID, 7 b 機器ID, 7 c 機種ID, 7 d 機器タイプID,
8 TCP/IPネットワーク, 9 WWWサーバ, 9 a HDD, 21 IEEE1394ネットワ
ークコントローラ, 21 a 機器接続検出部, 21 b, 22 音声UI定義情報変
換部, 22 a XMLパーサ, 22 b 音声認識辞書生成部, 23 音声認識部,
24 UI制御部, 24 a 音声認識制御部, 24 b UIコマンド決定部, 24
c UIコマンド履歴, 25 IEEE1394AV/Cコマンド変換部, 26 ROM, 27
ネットワークコントローラ, 28 ドライブ, 29 HDD, 30 音声UI定義
情報生成部, 31 無線通信部, 41 音声UI定義情報記憶部, 41 a乃至4
1 d 音声認識辞書, 101 磁気ディスク, 102 光ディスク, 103 光
磁気ディスク, 104 半導体メモリ

【書類名】 図面

【図1】

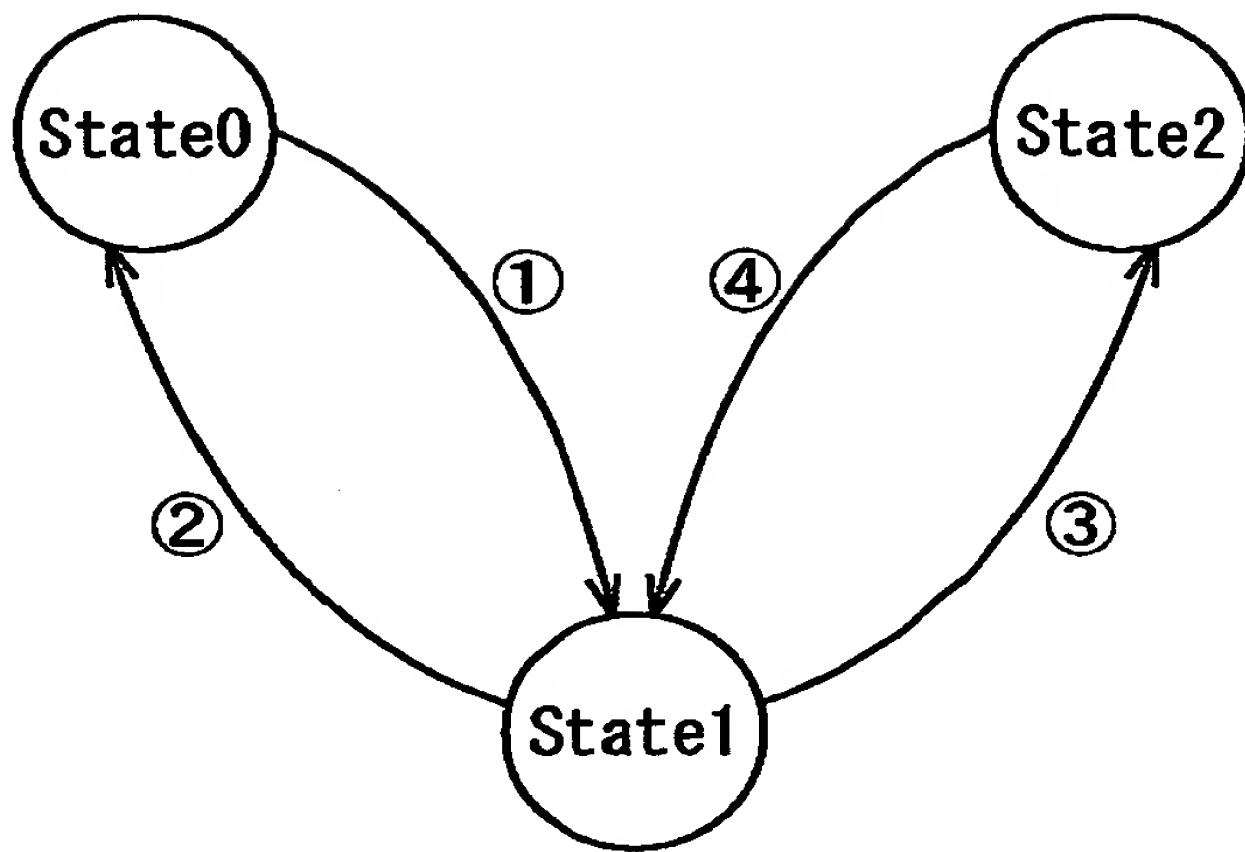


【図 2】

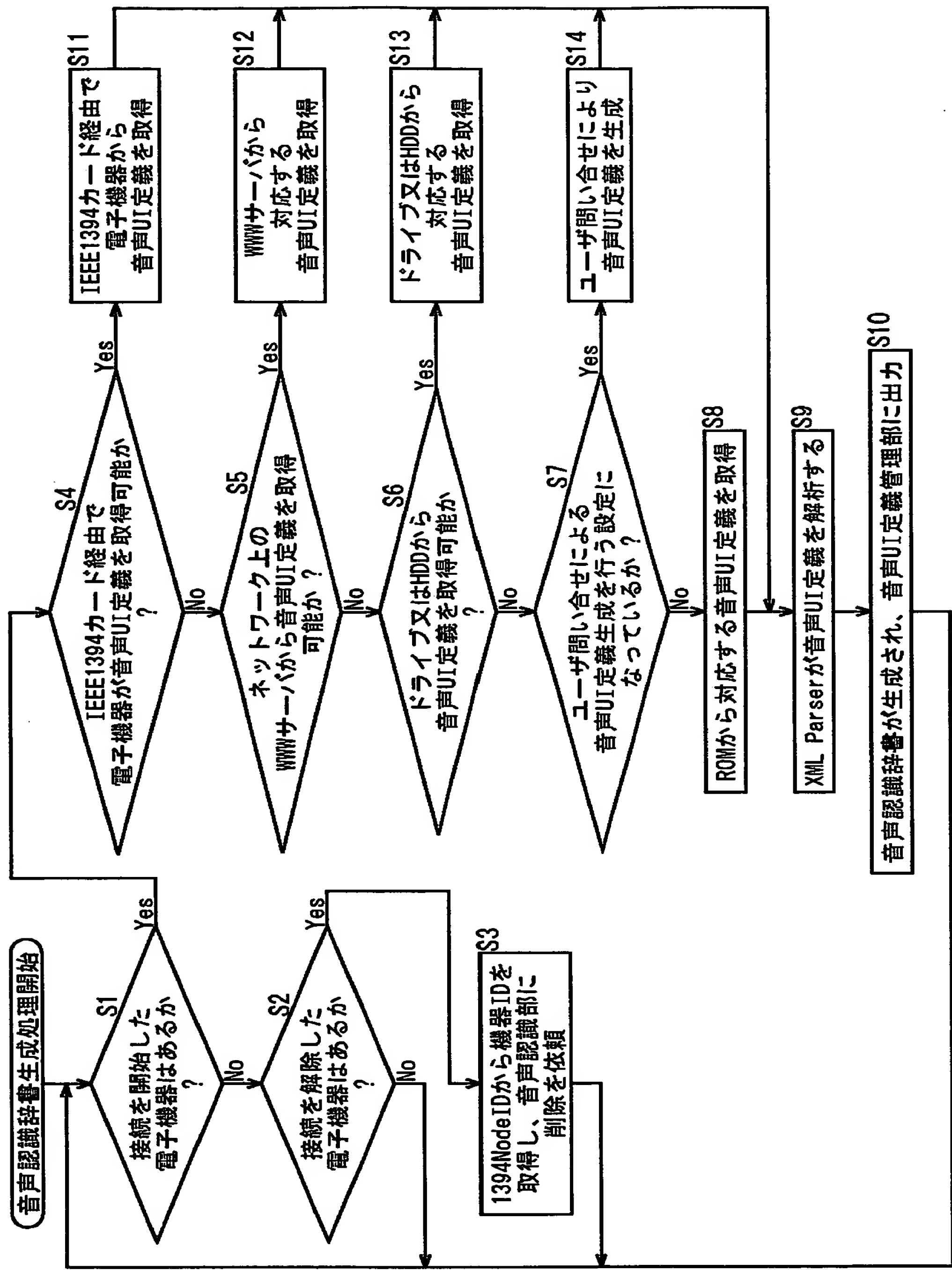
DVC

State 0				State 2			
音声コマンド	UIコマンド	遷移先State		音声コマンド	UIコマンド	遷移先State	
電源オン	Power On	1		停止	Stop	1	
電源を入れて	Power On	1					
State 1				...			
音声コマンド	UIコマンド	遷移先State					
電源オフ	Power Off	0					
電源を切って	Power Off	0					
再生	Play	2					
State X							
音声コマンド	UIコマンド	遷移先State		音声コマンド	UIコマンド	遷移先State	

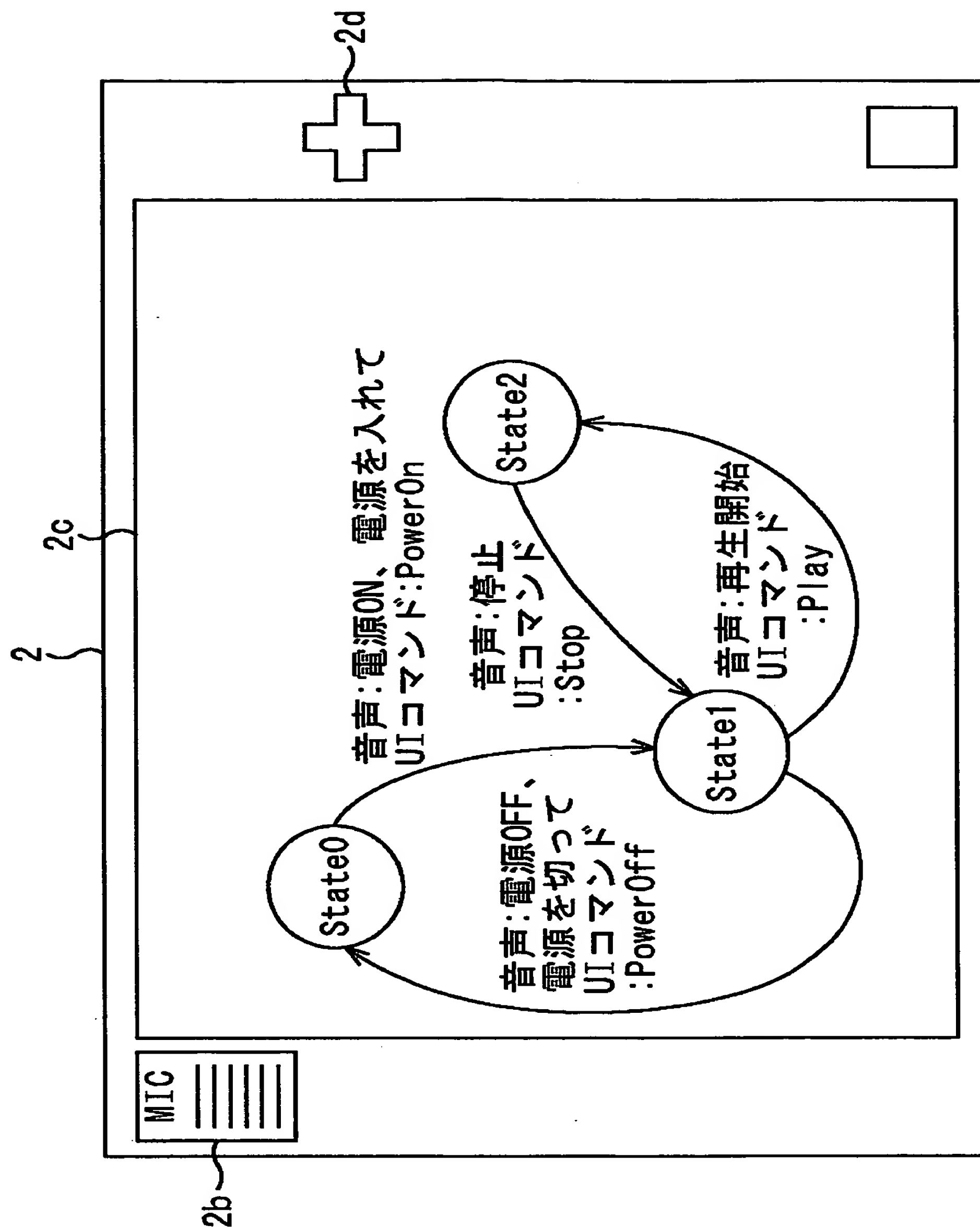
【図 3】



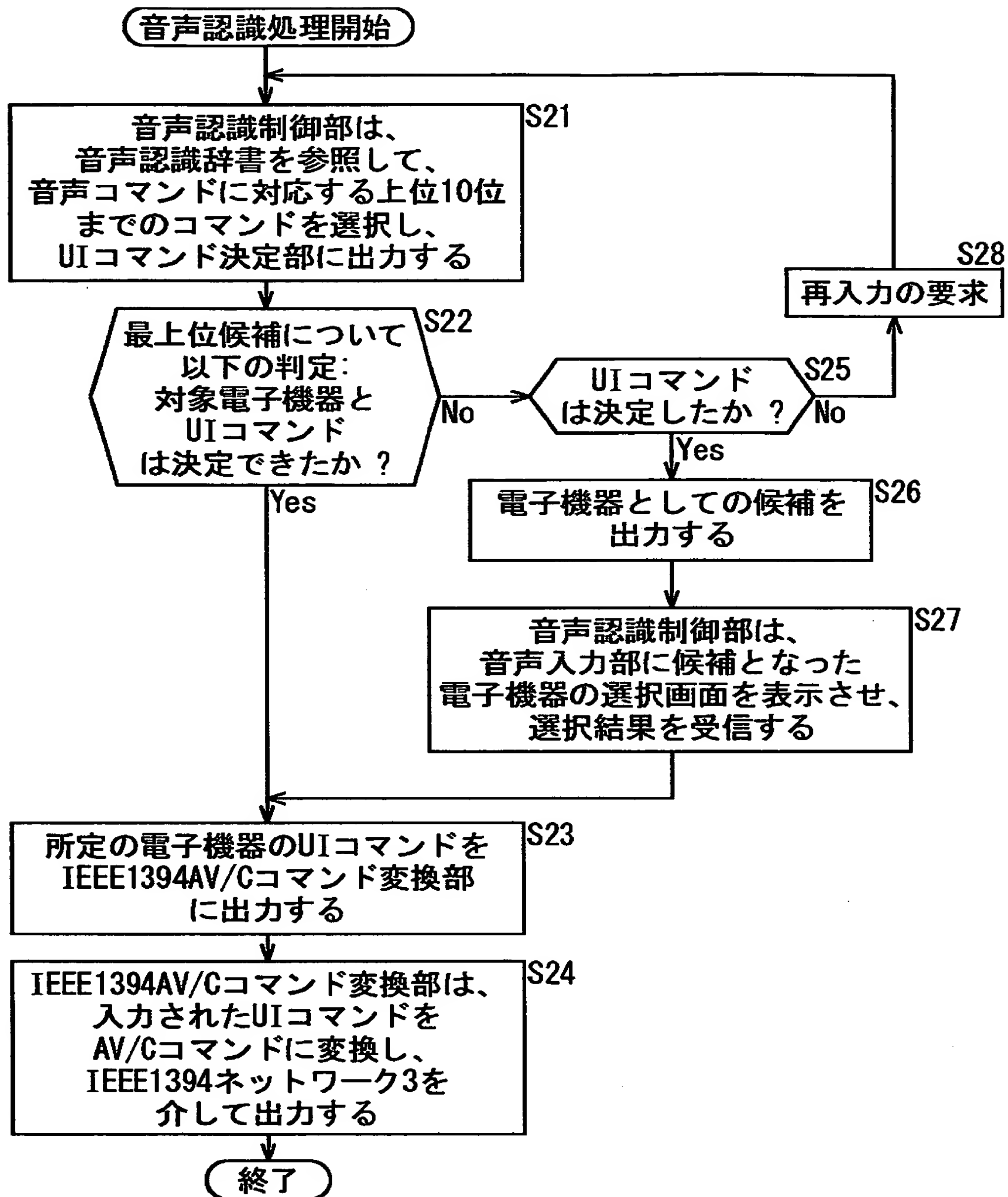
【図 4】



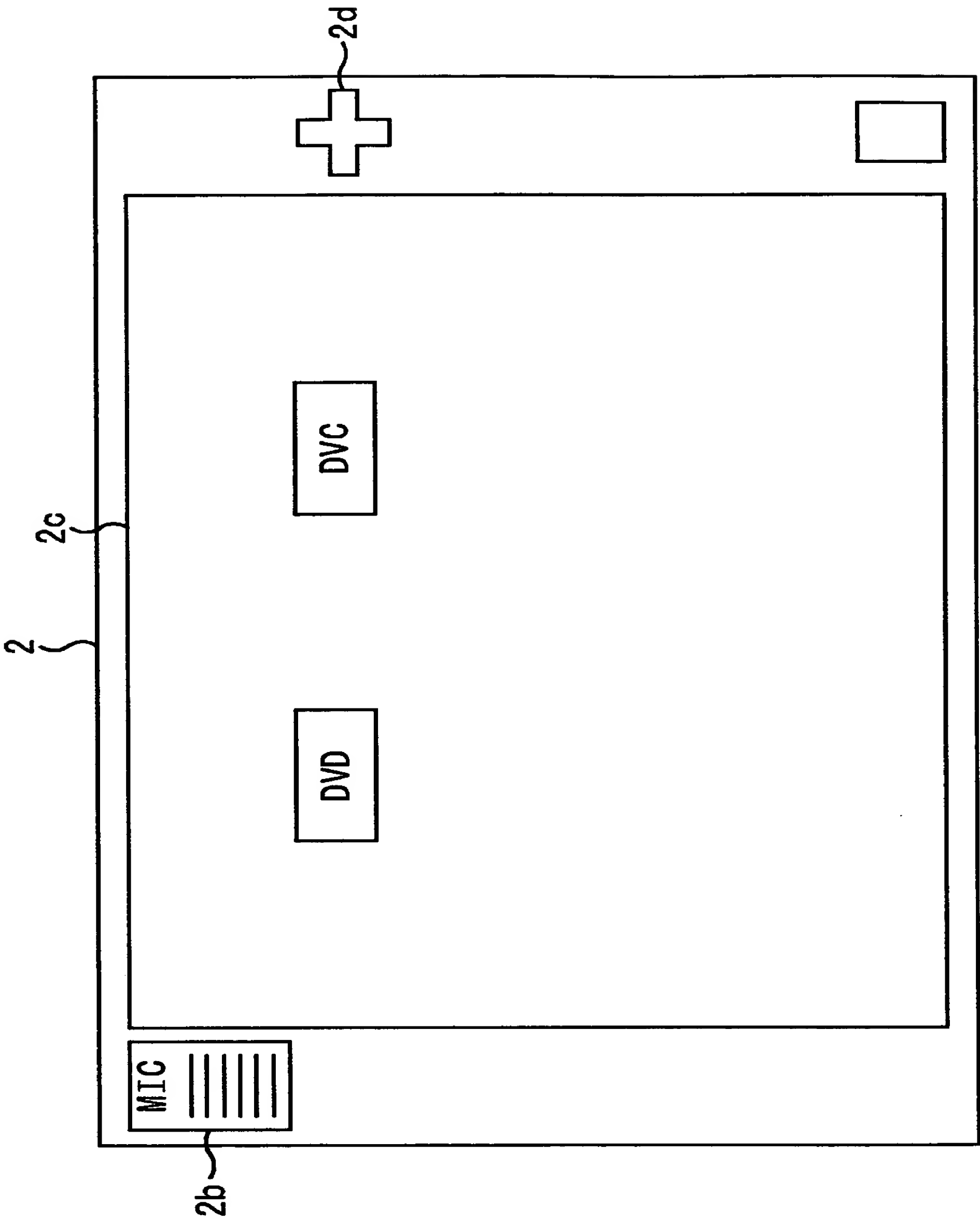
【図 5】



【図 6】



【図 7】



【書類名】 要約書

【要約】

【課題】 音声認識装置と音声入力インターフェースを統一して、音声コマンドによりネットワーク上の複数の電子機器を制御できるようにする。

【解決手段】 音声UI定義管理部 3 2 は、音声UI定義情報を取得し、XMLパーサ 2 2 a および音声認識辞書生成部 2 2 b を介して、音声UI定義情報記憶部 4 1 に音声認識辞書を記憶させる。音声認識制御部 2 4 a は、音声入力部 2 から入力される音声コマンドを、音声認識辞書を参照して認識し、候補となるUIコマンドをUIコマンド決定部 2 4 に出力する。UIコマンド決定部 2 4 は、UIコマンド履歴を参照し、所定のUIコマンドを選択して、IEEE1394AV/Cコマンド変換部 2 5 に出力する。IEEE1394AV/Cコマンド変換部 2 5 は、入力されたUIコマンドをAV/Cコマンドに変換し、所定の電子機器に出力する。

【選択図】 図 1

出 願 人 履 歴 情 報

識別番号 [000002185]

1. 変更年月日 1990年 8月30日
[変更理由] 新規登録
住 所 東京都品川区北品川6丁目7番35号
氏 名 ソニー株式会社